

## Interesse, Vernunft und Moral: zur sozialen Konstruktion von Vertrauen

Ziegler, Rolf

Veröffentlichungsversion / Published Version  
Sammelwerksbeitrag / collection article

### Empfohlene Zitierung / Suggested Citation:

Ziegler, R. (1997). Interesse, Vernunft und Moral: zur sozialen Konstruktion von Vertrauen. In S. Hradil (Hrsg.), *Differenz und Integration: die Zukunft moderner Gesellschaften ; Verhandlungen des 28. Kongresses der Deutschen Gesellschaft für Soziologie in Dresden 1996* (S. 241-254). Frankfurt am Main: Campus Verl. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-140111>

### Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

# Interesse, Vernunft und Moral: zur sozialen Konstruktion von Vertrauen

*Rolf Ziegler*

An einem nicht-alltäglichen, aber keineswegs unrealistischen Beispiel hat Thomas Schelling (1963: 43f.) einige Grundprobleme des Vertrauens in einer dyadischen Beziehung auf dramatische Weise veranschaulicht. Ein flüchtender Geiselnnehmer möchte sich seiner Geisel entledigen. Eigentlich möchte er sie freilassen, aber sie hat ihn erkannt. Sie beteuert, ihn nicht der Polizei zu verraten, und er möchte ihr gerne vertrauen, aber kann er das? Wie kann sie ihr Versprechen glaubhaft machen, wie kann sie sich glaubwürdig selbst binden? Gelingt es beiden nicht, ihr »Vertrauensproblem« zu lösen, wird der Geiselnnehmer die Geisel töten. Vertraut er dagegen der Geisel, dann muß er eine für ihn »riskante Vorleistung« erbringen, wie Niklas Luhmann (1973: 23) sagt.

Nun fehlt im Schellingschen Beispiel ein Aspekt: die »normative Absicherung« des Vertrauens. Niemand wird der Geisel einen moralischen Vorwurf machen, wenn sie später ihr Versprechen brechen sollte. Dagegen ist in Alltagssituationen häufig Vertrauen ein »normgeladenes« Phänomen, was bereits umgangssprachliche Ausdrücke wie Vertrauens»bruch« oder »Mißbrauch« des Vertrauens andeuten. Im Alltagsbewußtsein ist aber auch der Konflikt mit dem nackten Eigeninteresse stets präsent: Man kann Vertrauen aus»nutzen«. Daß jedoch nicht immer derjenige, der jemand fälschlicherweise vertraut, das Recht auf seiner Seite hat, zeigen nicht nur das Schellingsche Beispiel, sondern auch der Vorwurf, man sei zu vertrauensselig gewesen und habe unvernünftig gehandelt. Schließlich versucht man, durch »vertrauensbildende Maßnahmen« Vertrauen gezielt zu schaffen. Interesse, Vernunft und Moral scheinen also gerade beim Phänomen des Vertrauens in einem eigenartigen Spannungsverhältnis zu stehen.

Dieser Beitrag soll Ansätze zu einer Erklärung von spontaner Entstehung, Stabilisierung und Verlust von Vertrauen und seiner normativen Absicherung skizzieren, die sich aus einer Rational Choice-Perspektive ergeben. Er be-

schränkt sich auf Probleme des zwischenmenschlichen Vertrauens und klammert Fragen des sog. »Systemvertrauens« aus. Die Fragestellung soll in fünf Problemfelder aufgefächert werden: (1) die Struktur von Situationen, die einen »Bedarf« an Vertrauen erzeugen; (2) spontane Vertrauensbildung unter eigeninteressierten Akteuren; (3) die Bedeutung kontingenter Erwartungen für die Entwicklung stabilen Vertrauens; (4) die spontane Entstehung und Stabilisierung eines Sanktionssystems zur Kontrolle von Vertrauensmißbrauch; und (5) das Problem des »obligatorischen« Charakters von Vertrauen.

## 1. Die Struktur von Situationen, die einen »Bedarf« an Vertrauen erzeugen

Was sind die Grundelemente einer »normgenerierenden« Situation (Ullmann-Margalit 1977), die für zwei Akteure ein »Vertrauensproblem« enthält? (1) Beide ziehen es vor, daß jeweils der Andere sich kooperativ verhält, d.h. vertraut bzw. sich als vertrauenswürdig erweist. (2) Der Vertrauende erleidet einen Schaden, wenn sein Vertrauen mißbraucht wird. (3) Häufig hat die Vertrauensperson einen Anreiz, das Vertrauen zu mißbrauchen. (4) In der Regel würden es beide vorziehen, daß beide sich kooperativ anstatt daß beide sich unkooperativ verhalten. (5) Häufig muß der Vertrauende den ersten für ihn irreversiblen Schritt tun – wie im Schellingschen Beispiel der Geiselnnehmer. Jedoch sind keineswegs alle Vertrauenssituationen in dieser Weise asymmetrisch. Das bekannte Gefangenendilemma erfordert von beiden Akteuren »Vertrauen« im Sinne einer riskanten Vorleistung. Der Staatsanwalt konfrontiert ja die beiden Gefangenen in getrennten Verhören mit dem Angebot, demjenigen die Milde der Kronzeugenregelung zukommen zu lassen, der gesteht, und der Drohung, beide wegen einer geringeren Straftat überführen zu können bzw. denjenigen, der nicht singt, aber von seinem Komplizen verraten wird, die volle Härte des Gesetzes spüren zu lassen. Obwohl es einige Varianten gibt, sind dies die beiden typischen zwischenmenschlichen Vertrauenssituationen: das symmetrische Gefangenendilemma, das beide mit dem Problem des Vertrauens und der glaubwürdigen Selbstbindung konfrontiert, und die asymmetrische Situation, in der ein Akteur sich entscheiden muß zu vertrauen, und der andere das Problem hat, sein Versprechen (Lahno 1995) glaubhaft zu machen.

Ist der homo oeconomicus instande, sein Vertrauensproblem zu lösen, und wenn ja wie? Um sofort einem Mißverständnis vorzubeugen: Es geht hier

um eine theoretische und keine empirische Frage, denn natürlich wissen wir alle, daß Menschen – zwar keineswegs immer aber doch häufig – Vertrauen schenken und sich dessen als würdig erweisen. Wir Sozialwissenschaftler sind als die Alfred Schütz'schen »Direktoren dieses Marionettentheaters« (1972: 48) gefordert zu zeigen, wie die von uns konstruierten »Puppen« auf der von uns gebauten »Bühne« nach dem von uns geschriebenen »Drehbuch« das Vertrauens»spiel« spielen werden. Im symmetrischen Gefangenendilemma ist das Ergebnis bekannt: Jeder Akteur besitzt eine dominante und zwar die unkooperative Strategie. Er braucht überhaupt nicht zu wissen, wie sich der andere tatsächlich verhält, denn er stellt sich immer besser, wenn er den anderen verpfeift, gleichgültig ob dieser ebenfalls singt oder dicht hält. Beim asymmetrischen Vertrauensproblem hat nur die Vertrauensperson (in der Regel) eine dominante Strategie, nämlich das Vertrauen auszunutzen. Der Vertrauende besitzt keine solche. Da er sich jedoch auf den Standpunkt des anderen stellen und die Situation aus dessen Perspektive betrachten kann (diese strategische Fähigkeit besitzt natürlich der homo oeconomicus!), erwartet er dieses Verhalten und wird daher die für ihn unter diesen Umständen (relativ) günstigere Alternative wählen, nämlich nicht vertrauen. Das Ergebnis ist offensichtlich für die beiden homines oeconomici fatal: Ihr individuell rationales Verhalten führt zu einem Resultat, das beide schlechter stellt, als wenn sie sich kooperativ verhalten hätten. Es trifft aber auch den Direktor des »Rational Choice-Marionettentheaters«: Seine superrationalen Puppen fallen sehenden Auges und absichtlich in die soziale Falle. Nun kann der Direktor, d.h. der spieltheoretisch modellierende Rational Choice-Theoretiker, auf einen m.E. nicht zu unterschätzenden Vorteil seines »Drehbuches« hinweisen: man weiß wenigstens, was gespielt wird; oder weniger metaphorisch ausgedrückt: Die spieltheoretische Konzeptualisierung ermöglicht eine präzisere Explikation der theoretischen Fragestellung und eine genauere begriffliche Differenzierung z.B. auch des Rationalitätsbegriffes, d.h. sie zwingt sie zu einer expliziten Spezifizierung der grundlegenden Annahmen und liefert einen Beweis der logischen Implikationen.

## 2. Spontane Vertrauensbildung unter eigeninteressierten Akteuren

Gemeinsames Merkmal aller hier zu untersuchenden Ansätze zur Erklärung von spontaner Entstehung, Stabilisierung und Verlust von Vertrauen ist die grundlegende Tatsache, daß dies für bestimmte (Gruppen von) Akteure(n) er-

wünschte Konsequenzen hat. Sie haben wie James Buchanan (Brennan/Buchanan 1985; Vanberg/Buchanan 1988) sagt, ein »konstitutionelles Interesse« an einem funktionierenden Vertrauen. Vielfach wird jedoch bei dieser quasi-funktionalistischen Argumentationsweise der Prozeßmechanismus nicht expliziert, durch den effizienzsteigerndes Vertrauen entsteht, aufrecht erhalten wird oder verloren geht. Es genügt jedoch nicht, darauf hinzuweisen, daß eine bestimmte institutionelle Regelung geeignet ist, eine problematische Situation zu verbessern bzw. zu entschärfen. Die entscheidende Aufgabe besteht vielmehr darin zu zeigen, welche sozialen Bedingungen die einzelnen Akteure motivieren, sich an der Schaffung und Aufrechterhaltung einer institutionellen Regelung zu beteiligen (Voss 1985).

Ein auch in der Geschichte bedeutsamer Mechanismus ist der Einsatz von Pfändern oder Geiseln (Raub/Keren 1993; Weesie/Raub 1996). Sie wirken sowohl über ihren schadenskompensierenden Effekt für den Vertrauenden als auch durch die den Anreiz zum Vertrauensbruch mindernde, selbst auferlegte, potentielle »Bestrafung« der Vertrauensperson. Diese Wirkung können Pfänder sogar in einer einmaligen Begegnung unter Fremden entfalten. Die meisten Ansätze gehen jedoch von wiederholbaren Situationen aus, in denen das Luhmannsche »Gesetz des Wiedersehens« gilt. Sind die Partner in dauerhafte Beziehungen eingebunden, dann kann ein hinreichend großer »Schatten der Zukunft« (Axelrod 1987), in dem auch zukünftige Gewinne und Verluste berücksichtigt werden, kooperative (in diesem Sinne »vertrauensvolle«) Verhaltensregelmäßigkeiten unter eigeninteressierten Akteuren entstehen lassen. Verfügen die Akteure über eine Hirschmansche »Exit-Option« (1974) und brechen die Interaktion bei einem Vertrauensmißbrauch ab, dann sind dem dauerhaften Erfolg von renitenten Schwindlern enge Grenzen gesetzt (Schüßler 1990). Darüber hinaus schafft das Problem der (Re)identifikation von »schwarzen Schafen« – der Verbreitung einer schlechten Reputation – einen Anreiz zur Bildung eines effizienten Signal- und Kommunikationssystems, durch das Dritte davor gewarnt werden, eine Vertrauensbeziehung mit einem Schwindler aufzunehmen. Die inhärente Problematik eines solchen Informationssystems – nämlich die Sicherung der Glaubwürdigkeit der Signalgeber und des Vertrauens der Signalempfänger – wird durch ein Kommunikationssystem, wie z.B. den Klatsch, bewältigt, das Gelegenheiten und Anreize für die kostengünstige Produktion eines öffentlichen Gutes, nämlich die Verbreitung des guten oder schlechten Rufes eines Akteurs, schafft und das über den individuellen Reputationsverlust hinaus die allgemein geteilte, abschreckende Erwartung bei potentiellen Schwindlern entstehen läßt, daß ihre Täuschungsversuche aufgedeckt werden (Ziegler 1990).

Neuere Arbeiten unter Verwendung der evolutionären Spieltheorie (Muel-ler 1990) versuchen ebenfalls, eine dynamische Perspektive in die Analyse der Normentstehung und Vertrauensbildung (Güth/Kliemt 1993) einzubringen. Dabei geht es nicht um eine inhaltliche Übernahme biologistischer Standpunkte im Sinne der genetischen Determination menschlichen Sozialverhaltens, sondern um eine Forschungsheuristik für die Konstruktion analoger Modelle der sozio-kulturellen Evolution. Das Analogon zum genetischen Vererbungsmechanismus sind bei einer individualistischen Theorie der Evolution von Regeln soziale Lernprozesse, bei denen das Individuum die Regel von einem Verhaltensvorbild übernimmt. Dieses soziale Lernverhalten entspricht der Konzeption begrenzter Rationalität und kann unter bestimmten Bedingungen zu optimaleren Ergebnissen führen als individuelles Lernen, insbesondere nämlich dann, wenn dieses sehr fehlerträchtig ist. Die evolutionäre Dynamik sozialer Lernprozesse kann also zu effizienten Ergebnissen führen. Allerdings sind die Resultate von der Ausgangssituation und dem eingeschlagenen Entwicklungspfad abhängig, d.h. ohne Berücksichtigung der »Geschichte« eines sozialen Systems läßt sich seine Entwicklung nicht vorher-sagen. In der häufig als ahistorisch eingeschätzten Theorie rationalen Handelns erhalten so historische Konstellationen für die Erklärung sozialer Prozesse einen systematischen Platz weit jenseits paradigmenfremder bloßer ad hoc-Annahmen.

### 3. Die Bedeutung kontingenter Erwartungen für die Entwicklung stabilen Vertrauens

Wenn Vertrauen eine riskante Vorleistung ist, scheint es naheliegend, sie als eine Entscheidung unter Risiko aufzufassen. Dies ist in der Tat die vorherrschende Herangehensweise bei Vertretern des Rational Choice-Ansatzes (Coleman 1991/92). Das Risiko besteht danach in der Wahrscheinlichkeit, daß sich die Vertrauensperson als nicht vertrauenswürdig erweist. Deshalb wird ein rationaler Akteur nur dann die riskante Vorleistung erbringen, wenn der »erwartete« (d.h. mit der Wahrscheinlichkeit, es mit einer vertrauenswürdigen Person zu tun zu haben, gewichtete) Gewinn den erwarteten Verlust übersteigt, der mit der komplementären Wahrscheinlichkeit eines Vertrauensbruches eintritt. Diese Konzeptualisierung trägt jedoch einer Besonderheit sozialer Interaktionen nicht Rechnung, die unter dem Stichwort der »doppelten Kontingenz von Erwartungen« in der Literatur immer wieder thematisiert

worden ist. Meine Erwartung eines vertrauenswürdigen Verhaltens ist nämlich keine bloße Prognose einer bestimmten Verhaltensdisposition meines Gegenüber (Güth/Kliemt 1995: 39ff.), sondern stützt sich auf zweierlei: erstens darauf, daß dieser seiner Entscheidung (auch) seine Erwartungen hinsichtlich meines Verhaltens zugrunde legt (und nicht z.B. völlig unabhängig von seinen eigenen Prognosen unbedingt stets vertraut oder mißtraut) sowie zweitens darauf, daß seine Erwartungen auch eine Prognose meiner Erwartungen beinhalten (Heiner/Schmidtchen 1995). Modelliert man diese wechselseitigen Luhmannschen Erwartungserwartungen und bedingten Vertrauensstrategien, dann muß man zwei mögliche Irrtümer berücksichtigen, die den bekannten Fehlern I und II aus der statistischen Entscheidungstheorie entsprechen: Ich kann irrtümlicherweise einen Vertrauensbruch erwarten (Fehler I) oder ich kann irrtümlicherweise die Honorierung des Vertrauens vorhersagen (Fehler II). Begrenzte Rationalität bedeutet nun, daß Akteure keine perfekt reliablen Erwartungserwartungen besitzen: Sie können in der einen und/oder anderen Hinsicht irren. Außerdem kann dieser Irrtum sowohl auf Ego's Fehleinschätzung von Alter's Strategie als auch auf Ego's Fehleinschätzung von Alter's Prognose hinsichtlich Ego's Verhalten beruhen. Dennoch ist es möglich, daß bei hinreichend günstiger Gewinn-Verlust-Relation und/oder hinreichend großer Reliabilität ihrer Prognosefähigkeit zwei rationale Akteure selbst bei einer einmaligen Begegnung sich wechselseitig vertrauen (Ziegler 1996). Es genügen also hinreichend reliable, kontingente Erwartungen (zunächst rein kognitiver Art!), um Vertrauensbeziehungen zu stabilisieren. Vertrauen ist in diesem Sinne – worauf schon Georg Simmel hinwies – ein »mittlerer Zustand zwischen Wissen und Nichtwissen« (1983: 263), d.h. zwischen perfekt reliabler und absolut unrelabler Prognosefähigkeit, wo man mit gleicher Wahrscheinlichkeit richtig oder falsch vorhersagt.

Die entscheidende Frage ist nun natürlich, unter welchen Umständen reliable, kontingente Erwartungen entstehen bzw. ihre Reliabilität zu- oder abnimmt. Die dabei ablaufenden Lernprozesse haben unterschiedliche Konsequenzen, je nachdem welcher Zurechnungsgrund für die mangelnde Reliabilität der Erwartungen diagnostiziert wird. Besteht er in der mangelnden eigenen oder fremden Prognosefähigkeit, dann kann sich diese allmählich verbessern oder umgekehrt allmählich verschlechtern. Insbesondere durch langsame Steigern dessen, was auf dem Spiele steht, gefolgt von Vertrauensserweisen, kann so ein »Vertrauensbildungsprozeß« eingeleitet werden, auf den in der Literatur immer wieder hingewiesen wird. Sieht man jedoch den Grund enttäuschter Erwartungen in der »Absicht« des anderen Akteurs (d.h. einer bestimmten Verhaltensstrategie), dann kommt es zur abrupten Korrektur der

eigenen Erwartungen und damit zu einem entsprechenden unmittelbaren Verhaltenswandel. Exkulpationsversuche oder Erklärungen für mangelndes Vertrauen pflegen deshalb auch stets auf Irrtümer (oder auf das zu erwartende Verhalten der Anderen) zu verweisen und eine eigene Absicht in Abrede zu stellen.

Hinreichend reliable Prognosefähigkeit, die durch hinreichend verlässliche Signale unterstützt wird, kann also zu kontingenten Erwartungen führen, die wechselseitiges, bedingtes Vertrauen als eine stabile Verhaltensregelmäßigkeit entstehen lassen. Sofern darüber hinaus die zuvor bereits beschriebenen Mechanismen (das Gewicht von Pfändern, der Schatten der Zukunft, das Bestehen einer Exit-Option oder der Schaden eines schlechten Rufes) den Anreiz zu unkooperativem Verhalten mindern, erhöht sich die Chance, daß selbst eine mittlere Prognosefähigkeit ausreicht, um wechselseitiges Vertrauen zu stabilisieren.

#### 4. Die spontane Entstehung eines Sanktionssystems

Die bisherigen Ansätze haben Vertrauen durch stabile, sich selbst erfüllende, reziproke Erwartungen erklärt, aber keine spezifischen, direkten Sanktionen vorausgesetzt, wie sie gemeinhin mit dem Normbegriff verbunden werden. Die Anreiz- und »Sanktions«mechanismen waren der normgenerierenden Situation inhärent, d.h. die »Sanktion« gegen den Vertrauensbrecher bestand in genau demselben »abweichenden«, d.h. nicht-kooperativen Verhalten der betroffenen potentiellen Normbenefiziäre oder im Abbruch der sozialen Beziehung bzw. dem Ausschluß des Abweichenden aus dem sozialen System. Reputationsverlust als solcher ist zunächst keine spezifische, direkte Sanktion; er wirkt nicht über den Entzug sozialer Wertschätzung – einem Heiratsschwindler dürfte dies ziemlich gleichgültig sein –, sondern wegen des drohenden Ausschlusses von ausbeutbaren sozialen Beziehungen.

Eine Analyse der strategischen Interdependenzstruktur, die durch die Erweiterung der Handlungsoptionen um die Alternative »spezifische, direkte Sanktionen« entsteht, läßt jedoch rasch ein Sanktionsdilemma erkennen, das ein Glaubwürdigkeits- und ein Trittbrettfahrerproblem zweiter Ordnung enthält. Im Fall einer einmaligen Begegnung mag die Androhung der Sanktionierung eines Vertrauensbruches nämlich nicht glaubwürdig erscheinen, weil sie mit Kosten für den Sanktionsgeber verbunden ist, die er im Ernstfall nicht zu tragen bereit ist. Kann eine Sanktion von mehreren ausgeübt werden, so ent-



steht eine nicht minder problematische Situation. Eine (erfolgreiche) Sanktion schafft nämlich für die Normbenefiziäre das kollektive Gut »Konformität« (d.h. kein Vertrauensbruch), ist aber für den oder die Sanktionsgeber mit Kosten verbunden. Die Interdependenzstruktur unter den Sanktionsgebern ist in der Regel durch ein Freiwilligendilemma charakterisiert (Diekmann 1986): Im Extremfall kann jeder mit seinem eigenen Einsatz das kollektive Gut »Konformität« durch Sanktionierung des Abweichenden verschaffen, zieht es aber vor, einen anderen die Dreckarbeit machen zu lassen. Gesucht ist also ein Freiwilliger, der der Katze die Schelle umhängt. Angesichts dieser Problematik ist es schon erstaunlich, mit welcher theoretischen Naivität in der Soziologie abweichenden Verhaltens (übrigens der verschiedensten Couleur, z.B. auch bei den labeling-Theoretikern, die ja ihr besonderes Augenmerk auf den Sanktionierungs- bzw. Stigmatisierungsprozeß, die »societal reaction«, richten) die Frage »wer sanktioniert und welche Anreize und Motive bewegen ihn dazu?« entweder mit Stillschweigen übergangen oder mit der alles und nichts erklärenden Floskel erledigt wird, daß halt »die Gesellschaft« abweichendes Verhalten sanktioniere, und damit Punktum.

Ohne dies im einzelnen ausführen zu können, kann man im Prinzip die bereits erwähnten, erklärenden Mechanismen für die »Lösung« des Sanktionsdilemmas unter rationalen Egoisten heranziehen. Der Schatten der Zukunft und der Nutzen eines Rufes, sich nicht alles gefallen zu lassen, verändern die relativen Gewichte im Sanktionsdilemma zugunsten einer glaubwürdigen Sanktionsandrohung. Die kontingenten Erwartungen von Sanktionsgeber und Normadressat stabilisieren bei hinreichend reliabler Prognosefähigkeit ein beiderseitiges Verhaltensmuster, das Konformität sichert, ohne ständig die glaubwürdig angedrohte Sanktion tatsächlich ausführen zu müssen.

Einige Autoren haben die Sanktionsproblematik eine Stufe weiter verfolgt. In voll institutionalisierten Situationen ist das Sanktionsverhalten seinerseits normiert. Das bedeutet zum einen die Kontrolle von Sanktionsexzessen (z.B. Lynchjustiz), zum anderen die Unterstützung von Sanktionsgebern bei der Sanktionsausübung, aber auch die Sanktionierung von Sanktionsgebern wegen einer unterlassenen Sanktion. Computersimulationen demonstrieren die potentielle Bedeutung dieser sekundären Normierung für die Aufrechterhaltung sozialer Normen (Axelrod 1986). Es ist zu vermuten, daß die nachlassende Unterstützung des Sanktionsprozesses durch Dritte eine entscheidende erste Phase beim beginnenden Verfall sozialer Normen und der Erosion von Vertrauen ist.

## 5. Das Problem des »obligatorischen« Charakters von Vertrauen

Auch gegen die bis jetzt dargestellten Ansätze werden die Vertreter des normativistischen Paradigmas einwenden, daß die Akteure zwar (im Sinne des Regel-Utilitarismus) nicht auf fallweise Nutzenkalküle abstellen, aber nach wie vor den sog. »externen Standpunkt« (Kliemt 1985) einnehmen und Normen bestenfalls als (kollektive) »Klugheitsregeln« interpretieren, die gewissermaßen Odysseussche Selbstbindungsmechanismen gegen die Willensschwäche sind (Elster 1984). Die eigentliche »moralische Dimension« werde nach wie vor nicht gesehen. Diese liegt allerdings nicht im »altruistischen« Charakter einer Norm begründet (Gegenbeispiele sind die »moralische Verwerflichkeit des Selbstmordes« in manchen Gesellschaften oder die Diskussion um das »Recht auf ein würdiges Sterben«), sondern besteht vor allem in der Einnahme des »internen Standpunktes«, d.h. der Akzeptanz der Norm als eigenständigem Handlungsgrund ohne eine gesonderte Interessenkalkulation, oder um Emile Durkheim zu zitieren: »Eine Regel ist nämlich nicht nur eine gewohnheitsmäßige Form des Handelns, sie ist vor allem eine verpflichtende Form des Handelns, d.h., sie ist in bestimmtem Umfang der individuellen Willkür entzogen« (1988: 45). Die »moralische Dimension« bzw. die »Einnahme des internen Standpunktes« ist nun nicht einfach identisch mit der Übernahme »altruistischer Präferenzen« (Taylor 1987) bzw. einer »sozialen Orientierung« (McClintock/Liebrand 1988), zumal bei einer solchen Lösung in der Regel offen bleibt, unter welchen Umständen und in welchem Ausmaß Alter's Präferenzen Teil von Ego's Nutzenargumenten werden, und vor allem, wann dies normativ erwartet wird.

Nun scheint es auf den ersten Blick naheliegend, im Rahmen eines Rational Choice-Ansatzes dem Aspekt der Verinnerlichung von Normen dadurch Rechnung zu tragen, daß man die Nutzenfunktion schlicht um ein zusätzliches Argument »gutes/schlechtes Gewissen« erweitert und die Wahl der Akteure unter Einbeziehung dieser »internen Sanktionen« analysiert. Diese in der Literatur durchaus anzutreffende »Lösungs«strategie ist jedoch äußerst unbefriedigend, im Grunde völlig ad hoc und läßt eine ganze Reihe von Fragen offen. Erstens, bei welcher Art von Normen wird erwartet, daß man sich nicht nur aus bloßer Gewöhnung (»traditional«) oder aus Furcht vor Sanktionen (»zweckrational«) konform verhält, sondern daß man dies aus »der richtigen Gesinnung heraus« (»wertrational«) tut? Es mag zwar genügen zu sagen, daß man aus Gewöhnung oder aus Furcht vor einem Strafmandat ein Parkverbot beachtet, aber als Begründung dafür, daß man sein Kind nicht mißhandelt, reichen diese Gründe sicher nicht aus. Wenn eine Entführerin der »Landshut«

die damaligen Geiseln um Vergebung bittet und im selben Atemzug die Erschießung des Landshut-Piloten als »unnötig« bezeichnet, dann wird man ihr kaum »die richtige Gesinnung«, nämlich eine innerlich akzeptierte Verpflichtung des Tötungsverbots, attestieren wollen. Zweitens, wie kann man erklären, daß ein Akteur die Norm internalisiert, d.h. diese »moralischen Präferenzen« entwickelt? Es gibt Ansätze, die Entwicklung von Metapräferenzen (Sen 1982), die die Auswahl zwischen Präferenzen steuern, als Ergebnis von »natürlichen Neigungen« zu analysieren, die selbst nicht schon moralisch überformt sind (Hegselmann/Raub/Voss 1986). Sie berücksichtigen aber m.E. nicht hinreichend einen dritten Punkt, warum es nämlich gerade bei »moralischen« Normen nicht erlaubt und den Akteuren in der Regel gar nicht »möglich« ist, das normativ Gebotene gegen das Nützliche direkt abzuwägen?

Die letzte Frage bedarf einiger Erläuterungen. Vorschläge, interne Sanktionen lediglich als zusätzliche Nutzenargumente einzuführen, sind – wie bereits erwähnt – wenig überzeugend. Notwendig ist eine Erweiterung des Menschenbildes des Homo oeconomicus, die jedoch bei den schottischen Moralphilosophen bereits zu finden ist, wie ein etwas längeres Zitat aus Adam Smith' drittem Teil seines Werkes »Theory of Moral Sentiments« mit dem bezeichnenden Titel »Über die Grundlagen der Urteile, die wir über unsere eigenen Gefühle und unser eigenes Verhalten fällen, und über das Pflichtgefühl« belegt.

»In ganz gleicher Weise richtet sich unsere moralische Beurteilung zunächst auf Charakter und Verhalten anderer Leute und wir alle sind nur allzusehr geneigt, unser Augenmerk darauf zu richten, wie jeder von ihnen uns berührt. Aber wir erfahren bald, daß andere Leute mit ihren Urteilen über unseren Charakter und unser Verhalten ebenso freigebig sind. Wir werden nun ängstlich darauf bedacht sein, in Erfahrung zu bringen, inwiefern wir ihren Tadel oder ihren Beifall verdienen, und ob wir ihnen wirklich als so angenehme oder unangenehme Geschöpfe erscheinen mußten (sic! – Hervorhebung R.Z.), als welche sie uns hinstellen. Wir fangen deshalb an, unsere Affekte und unser Betragen zu prüfen und Betrachtungen darüber anzustellen, wie diese ihnen erscheinen müssen, indem wir bedenken, wie sie uns wohl erscheinen würden, wenn wir uns an ihrer Stelle befänden. Wir stellen uns uns selbst als die Zuschauer unseres eigenen Verhaltens vor und trachten nun, uns auszudenken, welche Wirkung es in diesem Lichte auf uns machen würde. Dies ist der einzige Spiegel, der es uns ermöglicht, die Schicklichkeit unseres Verhaltens einigermmaßen mit den Augen anderer Leute zu untersuchen« (1926: 169f.).

Bis in die Wahl der Spiegelmetapher hinein hat hier Adam Smith vor fast zweieinhalb Jahrhunderten die *topoi* des »role-taking«, »looking-glass self«

und (mit seinem »impartial spectator«) des »generalized other« vorweggenommen.

In der gebotenen Kürze soll – in Stichworten – die Argumentation skizziert werden. (1) Menschen streben – neben »physischem Wohlbefinden« – nach »sozialer Wertschätzung«. (2) Menschen gewähren soziale Wertschätzung denjenigen, die zur Verwirklichung ihrer »konstitutionellen Interessen« beitragen, und zwar (insbesondere) dann, wenn sie es »um der konstitutionellen Interessen willen, unbeschadet eigener Vor- und Nachteile« tun. Eine solche »bedingte« Gewährung sozialer Wertschätzung ist – falls erfolgreich – sehr effizient, da sie erhebliche Kontrollkosten spart. (3) Menschen sind – siehe das Zitat von Adam Smith – in der Lage, sich selbst mit den Augen der anderen zu sehen und die Motive der anderen bei der Gewährung der sozialen Wertschätzung nachzuvollziehen. (4) Der Mensch, sagt Adam Smith, »... will nicht nur Lob, sondern Lobenswürdigkeit, er wünscht so zu sein, daß er, wenn er auch von niemandem gelobt werden würde, dennoch der natürliche und schickliche Gegenstand des Lobes wäre. Er fürchtet nicht nur Tadel, sondern Tadelnswürdigkeit, er fürchtet so zu sein, daß er, wenn er auch von niemandem getadelt würde, dennoch der natürliche und schickliche Gegenstand des Tadels wäre« (1926: 172). (5) Als »würdig« – so die These – werden Menschen beurteilt und empfinden sich selbst, wenn sie die soziale Wertschätzung von ihren signifikanten Anderen »zu Recht« erhalten, das heißt, wenn sie »aus der richtigen Gesinnung« heraus gehandelt haben. (6) Das festzustellen, sind sie selbst am besten in der Lage. Ob dauerhafte Selbsttäuschung über die eigenen Motive möglich ist, sei hier offen gelassen. Für Soziologen wäre ein solches Phänomen gemäß der Methode der abnehmenden Abstraktion nur dann relevant, wenn es sich für die Erklärung kollektiver Explananda als systematisch notwendig erweisen sollte. (7) Dann und nur dann, wenn die Bedingungen der »Würdigkeit« erfüllt sind bzw. nicht vorliegen, vermag der Mensch sich selbst die soziale Wertschätzung als antizipierte, »innere Sanktion« zu erteilen bzw. sich ihrer zu entziehen.

Allerdings muß man der unvollständigen Internalisierung von Normen Rechnung tragen und den stets fortbestehenden Widerstreit zwischen einer strikt »normativen« Orientierung und einer rein »opportunistischen« Abwägung von Nutzen und Kosten (einschließlich der zu erwartenden direkten, externen Sanktionen) berücksichtigen. Dafür eignen sich »Framing-Modelle«, wie sie in jüngster Zeit entwickelt wurden (Lindenberg 1983, 1993; Esser 1991, 1996; Ziegler 1994). Die Theorie des Framing nimmt an, daß der Mensch aufgrund seiner begrenzten Rationalität oft zu einer Vereinfachung der Zielstruktur gezwungen ist, so daß er in einer Handlungssituation jeweils

(nur) ein dominantes Ziel direkt verfolgt. Dieses »Leitmotiv« strukturiert die Situation, während alternative Rahmen »im Hintergrund« bleiben und von dort aus die Dominanz des Leitmotivs verstärken oder abschwächen. Wendet man diese Idee auf normgeleitetes Handeln an, so muß man davon ausgehen, daß es (mindestens) zwei Rahmen gibt, einen »normativen«, bei dem das »konstitutionelle Interesse« der Normbenefiziere die Rangordnung der Ergebnisse bestimmt, und einen »opportunistischen« Rahmen, der durch die »natürlichen Neigungen« der Normadressaten definiert ist. Unter bestimmten Bedingungen kommt es zu einem »frame-switch«, d.h. der bisher latente Rahmen wird zum Leitmotiv und der zuvor dominante bildet nunmehr den Hintergrund.

## 6. Schlußbemerkung

In der Soziologie sind Versuche, normorientiertes Handeln als »rational« i.S. von nutzenmaximierend zu konzeptualisieren, seit jeher auf vehemente und ernst zu nehmende Kritik gestoßen. Die u.a. von Durkheim und Parsons vertretene normativistische Theorietradition sieht in den durch Sozialisation verinnerlichten, allgemein geteilten Normen und Werten selbständige, nicht auf die Eigeninteressen rückführbare Handlungsmotive, die im Gegenteil die nicht-kontraktuellen Voraussetzungen (interessegeleiteter) Verträge und Vereinbarungen sind. Eine sehr grundsätzliche Kritik – mit einer ebenfalls ehrwürdigen sozialphilosophischen Tradition (Nida-Rümelin 1993) – vertritt den Standpunkt, daß normorientiertes Handeln nicht zweckrational ist, weil es nicht an den Folgen, d.h. nicht an der Zukunft orientiert ist. Schließlich ist unter der Prämisse einer einzelfallorientierten Nutzenmaximierung der homo oeconomicus nicht imstande, sich glaubhaft selbst an Regeln zu binden (Güth/Kliemt 1995).

Am Beispiel des zwischenmenschlichen Vertrauens als einem Teil des umfassenderen Problems der sozialen Ordnung sollte gezeigt werden, wie ein an das von den schottischen Moralphilosophen entwickelte Paradigma anknüpfender Rational Choice-Ansatz unter Abschwächung des Rationalitätspostulates im Sinne einer »begrenzten Rationalität« und unter Berücksichtigung einer »dynamischen, evolutionären Perspektive von Interaktionssystemen« den theoretischen Hiatus zwischen normorientiertem und interessegeleitetem Handeln zu überbrücken versucht.

## Literatur

- Axelrod, Robert (1986), An Evolutionary Approach to Norms, in: *American Political Science Review* 80: 1095-1111.
- Axelrod, Robert (1987), *Die Evolution der Kooperation*. München, zuerst 1984.
- Brennan, Geoffrey/Buchanan, James M. (1985), *The Reason of Rules. Constitutional Political Economy*. Cambridge.
- Coleman, James S. (1991/92), *Grundlagen der Sozialtheorie*. 3 Bände. München, zuerst 1990.
- Diekmann, Andreas (1986), Volunteer's Dilemma. A Social Trap Without a Dominant Strategy and Some Empirical Results, in: Andreas Diekmann/Peter Mitter (Hrsg.), *Paradoxical Effects of Social Behavior*. Heidelberg/Wien.
- Durkheim, Emile (1988), *Über soziale Arbeitsteilung. Studie über die Organisation höherer Gesellschaften*. Frankfurt, zuerst 1893.
- Elster, Jon (1984), *Ulysses and the Sirens. Studies in Rationality and Irrationality*. Cambridge.
- Esser, Hartmut (1991), Alltagshandeln und Verstehen. Zum Verhältnis von erklärender und verstehender Soziologie am Beispiel von Alfred Schütz und »Rational Choice«. Tübingen.
- Esser, Hartmut (1996), Die Definition der Situation, in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 48: 1-34.
- Güth, Werner/Kliemt, Hartmut (1993), Menschliche Kooperation basierend auf Vorleistungen und Vertrauen. Eine evolutionstheoretische Betrachtung, in: *Jahrbuch für Neue Politische Ökonomie* 12: 253-277.
- Güth, Werner/Kliemt, Hartmut (1995), Elementare spieltheoretische Modelle sozialer Kooperation, in: *Ökonomie und Gesellschaft. Jahrbuch* 12: 12-62.
- Hegselmann, Rainer/Raub, Werner/Voss, Thomas (1986), Zur Entstehung der Moral aus natürlichen Neigungen. Eine spieltheoretische Spekulation, in: *Analyse & Kritik* 8: 150-177.
- Heiner, Ronald A./Schmidtchen, Dieter (1995), *Rational Cooperation in One-shot Simultaneous PD-Situations*. Fairfax, VA (mimeo).
- Hirschman, Albert O. (1974), *Abwanderung und Widerspruch*. Tübingen, zuerst 1970.
- Kliemt, Hartmut (1985), *Moralische Institutionen. Empiristische Theorien ihrer Evolution*. Freiburg/München.
- Lahno, Bernd (1995), *Versprechen. Überlegungen zu einer künstlichen Tugend*. München.
- Lindenberg, Siegwart (1983), Utility and Morality, in: *Kyklos* 36: 450-468.
- Lindenberg, Siegwart (1993), Framing, Empirical Evidence, and Applications, in: *Jahrbuch für Neue Politische Ökonomie* 12: 11-38.
- Luhmann, Niklas (1973), *Vertrauen. Ein Mechanismus der Reduktion sozialer Komplexität*. 2., erw. Auflage. Stuttgart.

- McClintock, Charles G./Liebrand, Wim B.G. (1988), Role of Interdependence Structure, Individual Value Orientation, and Another's Strategy in Social Decision Making: A Transformational Analysis, in: *Journal of Personality and Social Psychology* 55: 396-409.
- Mueller, Ulrich (1990) (Hrsg.), *Evolution und Spieltheorie*. München.
- Nida-Rümelin, Julian (1993), *Kritik des Konsequentialismus*. München.
- Raub, Werner/Keren, Gideon (1993), Hostages as a Commitment Device. A Game-Theoretic Model and an Empirical Test of Some Scenarios, in: *Journal of Economic Behavior and Organization* 21: 43-67.
- Schelling, Thomas C. (1963), *The Strategy of Conflict*. New York.
- Schüßler, Rudolf (1990), *Kooperation unter Egoisten: vier Dilemmata*. München.
- Schütz, Alfred (1972), *Gesammelte Aufsätze Bd. I: Das Problem der sozialen Welt*. Den Haag.
- Sen, Amartya (1982), *Choice, Welfare and Measurement*. Oxford.
- Simmel, Georg (1983), *Soziologie. Untersuchungen über die Formen der Vergesellschaftung*. 6. Auflage. Berlin, zuerst 1908.
- Smith, Adam (1926), *Theorie der ethischen Gefühle*. Leipzig, zuerst 1759.
- Taylor, Michael (1987), *The Possibility of Cooperation*. Cambridge.
- Ullmann-Margalit, Edna (1977), *The Emergence of Norms*. Oxford.
- Vanberg, Viktor/Buchanan, James M. (1988), Rational Choice and Moral Order, in: *Analyse & Kritik* 10: 138-160.
- Voss, Thomas (1985), *Rationale Akteure und soziale Institutionen. Beitrag zu einer endogenen Theorie des sozialen Tauschs*. München.
- Weesie, Jeroen/Raub, Werner (1996), Private Ordering: A Comparative Institutional Analysis of Hostage Games, in: *Journal of Mathematical Sociology* 21: 201-240.
- Ziegler, Rolf (1990), The Kula: Social Order, Barter, and Ceremonial Exchange, in: Michael Hechter/Karl-Dieter Opp/Reinhard Wippler (Hrsg.), *Social Institutions. Their Emergence, Maintenance and Effects*. New York.
- Ziegler, Rolf (1994), Normen, Rational Choice und Netzwerkanalyse, in: Hans-Ulrich Derlien/Uta Gerhardt/Fritz W. Scharpf (Hrsg.), *Systemrationalität und Partialinteresse*. Baden-Baden.
- Ziegler, Rolf (1996), *Social Norms and Rational Choice: Trust and Contingent Expectations*. München (mimeo).